CLAIMS

We claim:

1. An apparatus for switching packets, each packet having a header portion, an optional corresponding tail portion, and a class of service indicator, said apparatus comprising:

a pipelined switch comprising:

a plurality of packet header buffers (PHBs);

an equal plurality of PHB pointers, each said PHB pointer pointing to a corresponding PHB; and

an equal plurality of pipeline stage circuits connected in a sequence, comprising at least a first stage circuit and a last stage circuit, wherein:

each said stage circuit begins an operation substantially simultaneously with each other;

each said stage circuit passes data to a next stage circuit in said sequence when every said operation performed by all said stage circuits is completed;

said first stage circuit reads said header portion and stores said header portion in said corresponding PHB using said corresponding PHB pointer; and

said last stage circuit outputs a modified header portion; and

a receive buffer manager (RBM) comprising:

a joining circuit connected to said pipelined switch wherein said modified header portion and said corresponding tail portion are joined to form a modified packet;

a receive queue manager connected to said joining circuit that buffers said modified packet in a receive packet buffer and enqueues said modified packet using said class of service indicator and a plurality of receive queues; and

a dequeue circuit connected to said receive queue manager and said receive packet buffer, wherein said dequeue circuit uses said class of service indicator to dequeue said modified packet to a switch fabric.

2. The apparatus as recited in Claim 1, wherein said plurality of pipeline stage circuits further comprise:

a pre-process circuit connected to said first stage circuit, wherein said pre-process circuit uses a second said PHB pointer to record first data in said corresponding PHB;

a pointer lookup circuit connected to said pre-process circuit that compares said header portion to a first data structure and determines a leaf pointer using said second PHB pointer;

a table lookup circuit connected to said pointer lookup circuit that uses said leaf pointer to access one or more sets of linked data structures and to fetch second data, wherein said table lookup circuit uses a third said PHB pointer to record said second data in said corresponding PHB; and

a post-process circuit using said third PHB pointer and connected to said table lookup circuit, wherein said post-process circuit uses a fourth said PHB pointer to record third data in said corresponding PHB;

wherein said last pipeline stage circuit comprises a gather circuit connected to said post-process circuit, and wherein said gather circuit uses said fourth PHB pointer to assemble said modified header portion.

3. The apparatus as recited in Claim 1, further comprising:

an input device that receives said packet; and

a first buffer connected between said input device and said first stage circuit, wherein said first buffer buffers said header portion and said tail portion.

4. The apparatus as recited in Claim 3, further comprising a multiplexer interposed between said first buffer and said first stage circuit, wherein said multiplexer multiplexes said header portion and said corresponding tail portion together.

5. The apparatus as recited in Claim 1, further comprising a second buffer interposed between said last stage circuit and said joining circuit.

6. The apparatus as recited in Claim 5, further comprising a multiplexer interposed between said last stage circuit and said second buffer, wherein said multiplexer multiplexes said modified header portion and said corresponding tail portion together.

7. The apparatus as recited in Claim 1, wherein said receive packet buffer comprises buffers of different sizes.

8. The apparatus as recited in Claim 1, wherein said receive packet buffer comprises buffers of equal size.

9. The apparatus as recited in Claim 1, wherein said receive queue manager comprises a congestion avoidance circuit utilizing a status of each said receive queue.

10. The apparatus as recited in Claim 9, wherein said status comprises a measure of average queue depth.

11. The apparatus as recited in Claim 1, further comprising a transmit buffer manager (TBM), said TBM comprising:

a third buffer that receives one or more packets from said switch fabric;

a transmit queue manager connected to said third buffer that buffers each said packet in a transmit packet buffer and enqueues said packet using said class of service indicator and a plurality of transmit queues; and

a dequeue circuit connected to said transmit queue manager and said transmit packet buffer, wherein said dequeue circuit uses said class of service indicator to dequeue each said packet.

12. The apparatus as recited in Claim 11, wherein said transmit packet buffer comprises buffers of different sizes.

13. The apparatus as recited in Claim 11, wherein said transmit packet buffer comprises buffers of equal size.

14. The apparatus as recited in Claim 11, wherein said transmit queue manager comprises a congestion avoidance circuit utilizing a status of each said transmit queue.

15. The apparatus as recited in Claim 14, wherein said status comprises a measure of average queue depth.

16. The apparatus as recited in Claim 11, further comprising a transmit FIFO connected to an output of said dequeue circuit.

17. An apparatus for switching packets in a communications network device comprising:

a buffer that receives one or more packets, said packets comprising a class of service indicator;

a transmit queue manager connected to said buffer that buffers each said packet in a transmit packet buffer and enqueues said packet using said class of service indicator and a plurality of transmit queues; and

a dequeue circuit connected to said transmit queue manager and said transmit packet buffer, wherein said dequeue circuit uses said class of service indicator to dequeue each said packet.

18. The apparatus as recited in Claim 17, wherein said transmit packet buffer comprises buffers of different sizes.

19. The apparatus as recited in Claim 17, wherein said transmit packet buffer comprises buffers of equal size.

20. The apparatus as recited in Claim 17, wherein said transmit queue manager comprises a congestion avoidance circuit utilizing a status of each said transmit queue.

21. The apparatus as recited in Claim 20, wherein said status comprises a measure of average queue depth.

22. The apparatus as recited in Claim 17, further comprising a transmit FIFO connected to an output of said dequeue circuit.

23. A method of switching packets, which comprises:

receiving a packet, said packet having a header portion, an optional corresponding tail portion, and a class of service indicator;

switching said packet through a pipelined switch having a plurality of packet header buffers (PHBs), an equal plurality of PHB pointers wherein each said PHB pointer points to a corresponding PHB, and an equal plurality of pipeline stages connected in a sequence, comprising at least a first stage and a last stage, said switching further comprising:

beginning said sequence in each said stage substantially simultaneously with each other said stage;

passing data to a next stage circuit in said sequence when every said operation performed by all said stage circuits is completed;

reading and storing said header in said corresponding PHB using said corresponding PHB pointer; and

outputting a modified header portion; and

buffering said modified header portion in a receive buffer manager (RBM), said buffering further comprising:

joining said modified header portion and said corresponding tail portion to form a modified packet;

buffering and enqueuing said modified packet using said class of service indicator; and

dequeuing said modified packet using said class of service indicator.

24. The method of Claim 23, wherein said switching further comprises:

recording first data in said corresponding PHB using a second said PHB pointer;

comparing said header portion to a first data structure and determining a leaf pointer using said second PHB pointer;

fetching second data using said leaf pointer to access one or more sets of linked data structures and recording said second data in said corresponding PHB using a third said PHB pointer;

post-processing said header portion using said third PHB pointer and recording third data in said corresponding PHB using a fourth said PHB pointer; and

assembling said modified header portion using said fourth PHB pointer.

25. The method of Claim 23, wherein said switching further comprises buffering said header portion and said tail portion prior to said switching.

26. The method of Claim 25, wherein said switching further comprises multiplexing said header portion and said tail portion together prior to beginning said sequence.

27. The method of Claim 23, wherein said buffering further comprises buffering said header portion and said tail portion prior to said joining.

28. The method of Claim 27, wherein said buffering further comprises multiplexing said header portion and said tail portion together prior to said joining.

29. The method of Claim 23, further comprising:

receiving one or more packets;

buffering and enqueuing each said packet using said class of service indicator and a plurality of queues; and

dequeuing each said packet using said class of service indicator.

30. The method of Claim 29, wherein said buffering uses a packet buffer comprising buffers of different sizes.

31. The method of Claim 29, wherein said buffering uses a packet buffer comprising buffers of equal size.

32. The method of Claim 29, wherein said buffering further comprises avoiding congestion using a status of each said queue.

33. The method of Claim 32, wherein said status comprises a measure of average queue depth.

34. The method of Claim 29, wherein said dequeuing uses a transmit FIFO.

35. A method of switching packets, which comprises:

receiving one or more packets having a class of service indicator;

buffering and enqueuing each said packet using said class of service indicator and a plurality of queues; and

dequeuing each said packet using said class of service indicator.

36. The method of Claim 35, wherein said buffering uses a packet buffer comprising buffers of different sizes.

37. The method of Claim 35, wherein said buffering uses a packet buffer comprising buffers of equal size.

38. The method of Claim 35, wherein said buffering further comprises avoiding congestion using a status of each said queue.

39. The method of Claim 38, wherein said status comprises a measure of average queue depth.

40. The method of Claim 35, wherein said dequeuing uses a transmit FIFO.

41. A computer system for interfacing with a communications network, comprising computer instructions for:

receiving a packet, said packet having a header portion, an optional corresponding tail portion, and a class of service indicator;

switching said packet through a pipelined switch having a plurality of packet header buffers (PHBs), an equal plurality of PHB pointers wherein each said PHB pointer points to a corresponding PHB, and an equal plurality of pipeline stages connected in a sequence, comprising at least a first stage and a last stage, said switching further comprising:

beginning said sequence in each said stage substantially simultaneously with each other said stage;

passing data to a next stage circuit in said sequence when every said operation performed by all said stage circuits is completed;

reading and storing said header in said corresponding PHB using said corresponding PHB pointer; and

outputting a modified header portion; and

buffering said modified header portion in a receive buffer manager (RBM), said buffering further comprising:

joining said modified header portion and said corresponding tail portion to form a modified packet;

buffering and enqueuing said modified packet using said class of service indicator; and

dequeuing said modified packet using said class of service indicator.

42. The computer system of Claim 41, wherein said switching further comprises:

recording first data in said corresponding PHB using a second said PHB pointer;

comparing said header portion to a first data structure and determining a leaf pointer using said second PHB pointer;

fetching second data using said leaf pointer to access one or more sets of linked data structures and recording said second data in said corresponding PHB using a third said PHB pointer;

post-processing said header portion using said third PHB pointer and recording third data in said corresponding PHB using a fourth said PHB pointer; and

assembling said modified header portion using said fourth PHB pointer.

43. The computer system of Claim 41, further comprising:

receiving one or more packets;

buffering and enqueuing each said packet using said class of service indicator and a plurality of queues; and

dequeuing each said packet using said class of service indicator.

44. A computer readable storage medium, comprising computer instructions for:

receiving a packet, said packet having a header portion, an optional corresponding tail portion, and a class of service indicator;

switching said packet through a pipelined switch having a plurality of packet header buffers (PHBs), an equal plurality of PHB pointers wherein each said PHB pointer points to a corresponding PHB, and an equal

plurality of pipeline stages connected in a sequence, comprising at least a first stage and a last stage, said switching further comprising:

beginning said sequence in each said stage substantially simultaneously with each other said stage;

passing data to a next stage circuit in said sequence when every said operation performed by all said stage circuits is completed;

reading and storing said header in said corresponding PHB using said corresponding PHB pointer; and

outputting a modified header portion; and

buffering said modified header portion in a receive buffer manager (RBM), said buffering further comprising:

joining said modified header portion and said corresponding tail portion to form a modified packet;

buffering and enqueuing said modified packet using said class of service indicator; and

dequeuing said modified packet using said class of service indicator.

45. The computer readable storage medium of Claim 44, wherein said switching further comprises:

recording first data in said corresponding PHB using a second said PHB pointer;

comparing said header portion to a first data structure and determining a leaf pointer using said second PHB pointer;

fetching second data using said leaf pointer to access one or more sets of linked data structures and recording said second data in said corresponding PHB using a third said PHB pointer;

post-processing said header portion using said third PHB pointer and recording third data in said corresponding PHB using a fourth said PHB pointer; and

assembling said modified header portion using said fourth PHB pointer.

46. The computer readable storage medium of Claim 44, further comprising:

receiving one or more packets;

buffering and enqueuing each said packet using said class of service indicator and a plurality of queues; and

dequeuing each said packet using said class of service indicator.

47. A computer data signal embodied in a carrier wave, comprising computer instructions for:

receiving a packet, said packet having a header portion, an optional corresponding tail portion, and a class of service indicator;

switching said packet through a pipelined switch having a plurality of packet header buffers (PHBs), an equal plurality of PHB pointers wherein each said PHB pointer points to a corresponding PHB, and an equal plurality of pipeline stages connected in a sequence, comprising at least a first stage and a last stage, said switching further comprising:

beginning said sequence in each said stage substantially simultaneously with each other said stage;

passing data to a next stage circuit in said sequence when every said operation performed by all said stage circuits is completed;

reading and storing said header in said corresponding PHB using said corresponding PHB pointer; and

outputting a modified header portion; and

buffering said modified header portion in a receive buffer manager (RBM), said buffering further comprising:

joining said modified header portion and said corresponding tail portion to form a modified packet;

buffering and enqueuing said modified packet using said class of service indicator; and

dequeuing said modified packet using said class of service indicator.

48. The computer data signal of Claim 47, wherein said switching further comprises:

recording first data in said corresponding PHB using a second said PHB pointer;

comparing said header portion to a first data structure and determining a leaf pointer using said second PHB pointer;

fetching second data using said leaf pointer to access one or more sets of linked data structures and recording said second data in said corresponding PHB using a third said PHB pointer;

post-processing said header portion using said third PHB pointer and recording third data in said corresponding PHB using a fourth said PHB pointer; and

assembling said modified header portion using said fourth PHB pointer.

49. The computer data signal of Claim 47, further comprising:

receiving one or more packets;

buffering and enqueuing each said packet using said class of service indicator and a plurality of queues; and

dequeuing each said packet using said class of service indicator.

50. An apparatus for switching packets, said packets having a header portion, a tail portion, and a class of service indicator, comprising:

a pipelined switch comprising a plurality of stage circuits connected in a sequence wherein:

each said stage circuit begins an operation substantially simultaneously with each other said stage circuit;

each said stage circuit passes data when every said operation performed by all said stage circuits is completed;

a last stage circuit that outputs a modified header portion;

a queue manager connected to said last stage circuit, wherein said queue manager:

joins said modified header portion and said tail portion to form a modified packet; and

routes said modified packet using said class of service indicator.